DOCUMENT RESUME

ED 147 338 TM 006 649

AUTHOR Harnqvist, Kjell
TITLE Primary Mental Abilities at Collective and Individual Level. No. 60.
INSTITUTION Gothenburg Univ. (Sweden). Inst. of Education.
PUB DATE Oct 77
NOTE 67p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Ability Identification; *Cognitive Ability; *Cognitive Tests; Elementary Secondary Education; *Factor Analysis; Grouping Procedures; *Individual Differences; Performance Factors; *Predictor Variables; Statistical Analysis; Student Characteristics; Tables (Data); *Test Results
IDENTIFIERS Sweden

ABSTRACT

According to Cronbach's method of data analysis--which decomposes individual scores into separate components for different levels of aggregation--data collected for an investigation of the development of individual differences in ability test scores from grade 4 through grade 9 were grouped into school district, class, and individual components. Demographic characteristics of neighborhoods, self-selection of academic versus general programs, instructional emphasis, and irregularities in test administration were used to explain the test score variation between schools and districts. First-order factor analyses at the individual level showed the standard primary mental abilities pattern. Second-order analyses grouped the primary factors into one "power" and one "speed" factor. These same factors were also found in the analyses of class means; however, the processes behind the factors at the class level were likely to be different from those operating at the individual level. (Author/MV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Reports from

THE INSTITUTE OF EDUCATION

UNIVERSITY OF GOTEBORG

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kjell Härnqvist

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

PRIMARY MENTAL ABILITIES AT COLLECTIVE AND INDIVIDUAL LEVEL

Kjell Härnqvist

No 60 October 1977

EDUCATION, UNIVERSITY OF GOTEBORG Box S-431 20 MÖLNDAL, SWEDEN

PRIMARY MENTAL ABILITIES AT COLLECTIVE AND INDIVIDUAL LEVEL

Kjell Härnqvist

Göteborg University, Sweden

This reanalysis of an earlier study was done at the Center for Advanced Study in the Behavioral Sciences, Stanford, where the author was a Spencer Fellow during the academic year 1976-77. The author wishes to thank Lee J. Cronbach for comments on a plan for the analysis and on some preliminary results; Jan-Eric Gustafsson and Sten Stureson, Göteborg University, and Douglas Danforth and Carol Treanor at the Center for data processing; Joan Warmbrunn at the Center for preparing the manuscript.

Primary Mental Abilities at Collective and Individual Level

Abstract

Scores in ability tests administered to pupils in grades 4-9 were decomposed in district, class and individual components. Demographic characteristics of neighborhoods, self-selection between academic and general programs, emphasis in instruction, and irregularities in test administration were used to explain the variation between schools and districts. First-order factor analyses at individual level showed the usual PMA pattern. Second-order analyses grouped the primary factors into one "power" and one "speed" factor. This latter structure was found also in the analyses of class means although the processes behind the factors were likely to be different from those operating at the individual level.

Primary Mental Abilities at Collective and Individual Level

In a recent paper, Cronbach (Note 1) has shown numerous important consequences of the choice of level of analysis for data from classroom research. The choice of overall, within, or between group analysis for regressions, covariance adjustments of treatment group means, multivariate relations, etc., may lead to dramatically different results. Cronbach recommends methods of analysis that decompose the individual scores in components for different levels of aggregation and result in separate estimates for different levels.

This paper presents a secondary analysis of test scores decomposed in district, class and individual components. The data were originally collected for an investigation of the development of individual differences in ability test scores from grade 4 through grade 9 (Härnqvist, 1960). A reanalysis (Härnqvist, 1973) studied the stability of profiles on test and factor level within and between occasions. Both studies used the individual scores as if there were no effects of grouping in classes or school districts.

Data and setting

The samples to be used in this analysis are shown in Table 1. They in Sweden comprise all classes in six grades in five comprehensive school districts, chosen because of their coverage of practically the whole school population in their areas. This was expected to keep demographic differences out of the differences between grades. On average there were around

480 pupils in each grade, roughly equally divided between boys and girls, and distributed between from 20 to 29 classrooms. About 90 percent of the pupils were of normal age for their grade, i.e., eleven in grade 4, sixteen in grade 9 as the tests were administered during spring term.

Insert Table 1 about here

For these pupils scores are available in twelve tests. Ten of these were chosen as markers of five Primary Mental Abilities factors. The remaining two tests seem to some extent to rely on manual speed and dexterity. Pattern Copying also resembles one of the markers of Flexibility of Closure in French, Ekstrom and Price (1963).

Verbal (V)	Synonyms	40 items
	Opposites	40
Reasoning (R)	Letter Grouping	30
	Figure Series	30
Spatial (S)	Metal Folding	30
	Block Counting	40
Numerical (N)	Addition	2 x 48
	Multiplication	2 x 48
Perceptual speed (P)	Identical Number	2 x 45
	Highest Number	2 x 90
Not categorized	Pattern Copying	2 x 10
	Mark Making	2 x 300

Within school districts the pupils had been assigned to classes in different manners for grades 4-6 and 7-9. In the elementary grades the

assignment was normally done before grade 1 according to a neighborhood principle, and demographic differences between neighborhoods, for instance rural vs. more densely populated areas of the communities, may have had an influence. In grades 7 and 8 the pupils could choose a second foreign language and more academic courses in mathematics and some other subjects, or they could just take a general curriculum. This differentiation was further developed in grade 9 where the choice was between an academic, a general and a vocational line with different branches. In grades 7 through 9 pupils with similar choices were brought together in classes in one central school unit for the whole district. Classes in grades 4-6 were taught by classroom teachers. Classes in grades 7-9 were taught by subject-matter teachers.

Methods of analysis

Three types of analyses will be presented. In the first step the overall variance in the scores of each test will be divided into three components:

variance between districts (D),

variance between classes within districts (CwD),

variance between pupils within classes (PwC).

These variances will be expressed as percentages of the overall variance of the test. Computationally this is done by means of transforming each pupil score into three components: his district mean, the deviation of his class mean from his district mean, and the deviation of his raw score from his class mean. Variances are then computed for the three

parts separately with the number of pupils in the grade as N. This means that the estimates at all three levels are weighted in relation to the number of pupils and can be added to get the overall variance (see Cronbach, op. cit., Ch. 3).

The proportions of overall variance ascribed to the district and class levels correspond to correlation ratios calculated from sums of squares in analysis of variance (Winer, 1971, p. 118):

$$e^2 = 1 - \frac{SS_{within}}{SS_{total}}$$

This is one among possible measures of the variability between groups. One alternative to this is correlation ratios calculated from mean squares (ibid.) which can be written

$$\hat{\eta}^2 = 1 - \frac{N - 1}{N - k} \cdot \frac{SS_{within}}{SS_{total}} .$$

This gives a lower and less biased estimate of the variability between groups. The ratio between degrees of freedom grows when the number of subgroups (k) increases and when the number of individuals per group (n) decreases. In the present data the ratio is 1.01-1.02 for districts, varies between 1.04 and 1.08 for intact classes and between 1.07 and 1.14 for classes when boys and girls are taken separately.

A second alternative measure is the intraclass correlation coefficient, defined by

$$\rho = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_\varepsilon^2}$$ (Winer, op. cit., p. 244)

where σ_τ^2 and σ_ε^2 are variance components for treatments and error. One estimate of the intraclass correlation is

$$\hat{\rho} = \frac{MS_{between} - MS_{within}}{MS_{between} + (n - 1)MS_{within}} \quad \text{(ibid., p. 258).}$$

From the observed variation between groups is subtracted the variation that can be expected from random assignment to groups of size n from a population with variance equal to the pooled within group variance. In the present case it does not matter whether the collectives to which the random assignment is made are regarded as random or fixed in the variance-component model.

No simple algebraic relation exists between the correlation ratios and the intraclass correlation, but a number of examples calculated from the present data indicate that $\hat{\rho}$ comes very close to $\hat{\eta}^2$.

In spite of the statistically attractive properties of $\hat{\eta}^2$ and $\hat{\rho}$, proportions of observed variance, corresponding to e^2 and $1 - e^2$, have been used in the present study for two reasons.

1) The decomposition in collective and individual levels will be done not only for the variance of each test as such but also for the covariance between tests, i.e., not only for the diagonal elements of the variance-covariance matrix but also for the off-diagonal elements. For those the decomposition cannot be based on other than observed covariation. In order to be consistent with that, the diagonal elements also should be based on observed variation without reduction for expected effects of random assignment.

2) Following Cronbach (op. cit., p. 4.2-4) one can argue that classroom collectives should be regarded as having a fixed membership and a common history, the effects of which are shown in the variability between collectives. Such collectives cannot be replicated in a random assignment sense. Then it is the full variation between collectives that should be analyzed and not only what remains after correcting for possible effects of random assignment. Such an argument seems to be more valid for what Cronbach (op. cit., p. 1.17 b-c) calls group-caused effects than for demographic effects where groups differ initially. In the latter case it may be meaningful to compare their variation also with a random assignment model. The distinction between these types of effects, however, is a matter of interpretation and this must be based also on other information than the size of variation. The choice of e^2 means that results reported here are descriptive for the set of districts, classes and pupils analyzed in this study. Generalization to other contexts cannot be based on formal statistical procedures.

The analysis of the off-diagonal elements together with the decomposed variances will be done by means of factor analysis at class and individual level. Because of the small number of districts, and the relatively small variation found between them in the first step, the district level will be omitted here and district variation merged with class variation so that each score now is decomposed in only two parts: class means and deviations of pupil scores from class means. Cronbach (op. cit., p. 9. 13) recommends not to use correlations as a basis for such two-level factor analysis but covariances brought to a common scale

by means of the overall variances of the tests. This procedure may deserve a more detailed description.

Covariance matrices are calculated at between-class and within-class levels. Since the differences in variance between tests to a large extent depend on the number of items, time limits and similar arbitrary conditions, it is desirable to eliminate such variation. This is done by means of dividing each covariance with the product of the standard deviations of the overall distributions for the pair of tests involved. The resulting "scale-free" covariances for between- and within-level add up to the correlation in the pooled data. This procedure thus divides the overall correlation matrix in two additive parts: one for classes and one for pupils within classes.

In the factor analyses these two parts of the overall correlations are treated separately and with the proportion of variance of each test at the between- and within-class level used as communalities. The between-class proportions correspond to correlation ratios (e^2), the within-class proportions are $1 - e^2$. The use of these proportions instead of communalities estimated from the matrix serves the function of bringing also the factor variances onto the common scale of the overall variance.

In principle this division of overall correlations in between- and within-group parts can be done also for reliability estimates. In the present data part scores for parallel forms administered on consecutive days are available for some tests, and the correlation of these part scores will be further analyzed. The reliability of the other

tests was determined from odd-even part scores which cannot be regarded as experimentally independent and so they will not be studied in this connection.

Results I: Variance proportions

The relative proportions of the total variance accounted for at district, class and individual level are shown in Table 2 for intact classes. In Appendix A corresponding tables are found for half-classes of boys and girls.

Insert Table 2 about here

The relative contributions to total variance from district, class and individual level differ between grades which is shown by the following averages.

	Average variation for all tests in grade					
	4	5	6	7	8	9
District	2	2	2	4	6	4
Classes within districts	12	9	15	20	22	28
Pupils within classes	86	88	82	76	72	69

At class level grades 7-9 have higher proportions than grades 4-6 and this was expected because of the different principles of class composition at the two stages. A closer analysis shows that the differences are largest for the following tests.

	Average class contributions in grades	
	4 - 6	7 - 9
Opposites	9	32
Letter Grouping	9	30
Synonyms	11	27
Figure Series	10	22
Multiplication	13	25
Addition	11	21

This list contains all tests that have been categorized as measuring verbal, reasoning and numerical abilities. In the spatial tests the main difference occurs between grades 8 and 9.

	Average class contributions in grades		
	4 - 6	7 - 8	9
Block Counting	7	12	27
Metal Folding	10	16	26

The remaining tests show average differences of a similar kind but they are smaller and less regular in relation to the school stages.

Thus the class contributions to variance in tests of a general intellectual and scholastic content change most at the point where self-selection between academic and non-academic programs takes place. When vocational programs -- some of them clearly practical-mechanical in their orientation -- are introduced in grade 9 class variability increases also in spatial tests.

Even district contributions to variance are higher in grades 7 - 9 than in grades 4 - 6. Here, however, the elective programs in the upper grades cannot be used for explaining the difference since self-selection

to programs takes place within districts and each district draws pupils from the same demographic background in all six grades.

An inspection of the larger district proportions may give some clues to interpretation. The following tests have two or more proportions of at least 5 percent.

	Grades with 5% or more
Mark Making	6, 7, 8, 9
Pattern Copying	4, 5, 6, 7
Multiplication	7, 8, 9
Addition	7, 8
Highest Number	7, 8
Synonyms	8, 9

A first observation is that this list includes five of the six tests where timing was expected to be important for test scores and where consequently two parallel forms were used in order to get reliability estimates. Testing times were between 3 and 5 minutes per form. With such short testing times variations in test taking sets, irregularities in timing and disturbances during administration can have considerable influence upon results.

The district means of the parallel forms of the "speed" tests have been inspected and deviations have been noted from the "normal" development from grade to grade and from the first to the second parallel form. In both comparisons considerable gains have been found in most instances but there are a number of striking exceptions.

In Multiplication and Addition, and similarly for both parallel forms, the means cease to grow or even decrease between grades 6 and 7 in all districts except one which continues to show gains of "normal" size. This difference is likely to be a treatment effect, due to differences between districts in emphasizing continued training of arithmetic skills. It explains the increase of district variance between the elementary grades and grade 7 -- a difference that remains in grade 8 and to some extent in grade 9 too.

In Highest Number and Mark Making several district means show irregular trends from the first to the second parallel form. These differences often can be traced back to extreme deviations from the usual pattern in individual classes which influence also the district means and contribute to district variance. Such differences seem to be caused by irregularities in test administration. To some extent this explanation holds also for Pattern Copying but there one can also find a few instances of systematic deviations of districts from the "normal" age trend which look like treatment effects.

These examples may suffice to show that the variations in district contributions can be traced back to specific sources of influence, also that the variance proportions are sensitive to such influences. A more complete interpretation, however, requires information about the teaching and testing conditions that are not available for these data.

So far the analysis has been based on Table 2 -- the results for intact classes with boys and girls together. The results for half-classes

with boys and girls taken separately (Appendix A) show similar trends but due to the smaller group sizes, quite naturally, more irregularities too. The within-class proportions are smaller for each sex taken separately than for the pooled data.

	Within-class proportions in grade					
	4	5	6	7	8	9
Observed for intact classes	86	88	82	76	72	69
Observed for half-classes	80	82	76	74	70	64

This difference can be explained in two principally different ways.

1) The exclusion of sex differences in the tests is likely to reduce the variances, and primarily those within the classes. If the variance contributions within intact classes were reduced, in proportion to the reduction of within-class variance when boys and girls are taken separately, the averages for half-classes would become:

	4	5	6	7	8	9
Expected for half-classes	76	80	74	72	68	63

These values are around two percent units lower than the observed half-class variation. The variance reduction thus explains even more than the observed difference which means that the exclusion of sex differences also must have affected the between-class proportions.

2) The measurement of variability by means of e^2 instead of $\hat{\eta}^2$ introduces a downward bias in the estimates of within-class proportions, but more so in the estimates for boys and girls separately than for intact classes (cf. previous section about ratios of degrees of freedom). This

can be tested by means of direct calculation of $1 - \hat{\eta}^2$.

		4	5	6	7	8	9
$1 - \eta^2$:	intact classes	93	94	88	78	75	73
	half-classes	91	93	86	79	75	72

This latter adjustment brings the two sets of proportions very close to each other. Thus it has been possible to explain the differences in observed within-class proportions twice and in principally different ways -- maybe too much of a good thing to be fully satisfactory! These latter estimates leave little room for group differences in the elementary grades but the change from grade 7 on is even more evident here.

Results II: Factor analyses

Examples of different measures of relations between tests are shown in Table 3.

Insert Table 3 about here

The first row contains the overall correlation, i.e., the type of measure commonly used in factor analysis. In the second and third row this correlation is decomposed into scale-free covariances at between- and within-class levels. These add up to the overall correlation. The last two rows show the correlations between class means and the pooled within-class correlations. The correlations between class means tend to become very high when the correlation ratios are of the size prevailing in grade 9.

Factor analyses have been performed on the scale-free covariances at within- and between-class levels. Only the ten tests that could be

categorized as markers of five primary mental abilities were included in the analyses. At between-class level two principal components were extracted, at within-class level five, corresponding to the number of hypothesized primary mental abilities. At both levels the components were rotated according to the direct quartimin method (BMDP, 1975, p. 368), which gives oblique factors. From the five correlated factors in the within-class analyses second order factors were extracted and rotated, also with the direct quartimin method.

Table 4 shows the distribution of the total variance of the battery between the unrotated factors in the data for intact classes.

Insert Table 4 about here

The subtotals between classes correspond to the average correlation ratios (e^2) expressed as percentages. These are higher in grades 7-9 than in grades 4-6 -- a finding that already has been interpreted in the previous section. The subtotals within classes show the opposite pattern. Almost all of these differences are accounted for by the first principal components within and between classes. The remaining parts of the variance are rather equally distributed between levels and components in all six grades. The components taken out for rotation explain between 81 and 83 percent of the battery variance.

The analyses for half-classes (see Appendix B) show similar trends but in these cases the between-class factors account for relatively more of the battery variance -- a finding that also has been discussed previously.

In the within-class analyses of classes with boys and girls together the same five factors appear in all grades except grade 4. The results are shown in Table 5. The factors have been reported in the same order of content in all grades but with the corresponding factor numbers also given.

Insert Table 5 about here

All analyses show a very neat simple structure and in grades 5 - 9 it corresponds almost exactly to that hypothesized in the composition of the test battery. Among the loadings above .24 only four -- all of them in the order of .30 -- appear in places where they were not expected. Grade 4, however, deviates from the usual PMA pattern in two respects. The reasoning and spatial tests go together in a different way and numerical and perceptual speed tests are not distinguished from each other.

In the analyses of half-classes consisting of boys and girls respectively (see Appendix C) the same regular pattern as in grades 5 - 9 above was found in grades 5 - 8 for boys and grades 6 - 9 for girls. Altogether in these eight instances there are eleven loadings in the order of .30 that appear in unexpected places. Grade 4 for boys and grades 4 - 5 for girls deviate from the main pattern in a similar way as grade 4 for intact classes. In grade 9 for boys reasoning and numerical tests go together in a most unexpected way.

The interpretation of the deviations will be postponed to the discussion. For the present it will be concluded that the majority of

instances at within-class level strongly support the usual categorization of the tests in five "primary mental abilities" -- V, R, S, N and P. This result corresponds to what Henrysson (1965) found in an analysis of the original pooled matrices.

Next the results of the second-order analyses for intact classes will be reported (Table 6).

Insert Table 6 about here

In grades 5 - 9 where the usual PMA pattern was found, the first-order factors R, S and V (in the order mentioned) form one second-order factor, and N and P (in varying order) another. The correlations between these two factors (Appendix F) vary between .14 and .32.

The second-order analyses for half-classes (see Appendix D) show a similar two-factor pattern in all instances where the PMA pattern was found at individual level. Also for half-classes the first second-order factor tends to be strongest in R.

The between-class analyses for intact classes are reported in Table 7.

Insert Table 7 about here

In grades 4 - 8 two distinct factors appear. One is loaded in tests belonging to factors V, R and S at within-class level, and one in N and P tests. This distinction corresponds to the main pattern found in the second-order analysis of within-class factors. It also corresponds to the

old distinction between "power" and "speed" tests. Grade 9 shows another pattern with the first factor comprising all tests and the second contrasting spatial against numerical and perceptual tests, mostly, however, with very small loadings. The correlations between the two factors (Appendix F) vary between .27 and .50 except in grade 9 where the orthogonal pattern was left unchanged. An inspection of the loadings, however, indicates that an alternative solution in better agreement with the other analyses could be found.

The separate analyses for boys and girls (see Appendix E) have the "power" vs. "speed" pattern in all grades among boys and in grades 4 - 7 among girls. In grade 8 for girls factor 1 is loaded in all tests except Highest Number, which appears in factor 2. In grade 9 for girls the verbal tests show up in both factors.

In conclusion it may be stated that second-order within-factors and first-order between-factors, in general, show the same pattern with verbal, reasoning and spatial abilities belonging to one factor and numerical and perceptual facilities to another.

Results III: Decomposition of part score correlations

The scores in Addition and Multiplication which together define a numerical factor in most within-class analyses are based on scores from two parallel forms each. So are the scores of the perceptual speed tests, Identical Numbers and Highest Number. The proportions of variances and covariances of these part scores will be used for analyzing the contributions of different sources of variation. Table 8, based on decomposed

variance-covariance matrices for all grades and for intact classes, summarizes these analyses for grades 4-6 and 7-9 separately.

Insert Table 8 about here

The variance proportions (rows a and d) are very similar to those found in Table 2 for the whole tests. The scale-free covariances within the tests (rows b and e) add up to the overall parallel test correlations expressed as percentages. Through application of the Spearman-Brown formula they can be translated into average reliability estimates, and these become around .90 for Addition and Multiplication and around .85 for Identical Numbers and Highest Number. The overall proportions in rows c, f and g correspond to average correlation coefficients between different tests.

From the data behind Table 8 more analytic indices can be calculated. Decomposed reliability coefficients are obtained from the ratios b/a and e/d inserted in the Spearman-Brown formula.

Reliability of		Grades 4-6	Grades 7-9
Numerical tests	Class level	.89	.94
	Pupil level	.90	.89
Perceptual tests	Class level	.78	.78
	Pupil level	.90	.89

The two tests have the same reliability at pupil level, i.e., within classes, but differ at class level. The lower coefficients for the perceptual speed tests may be interpreted as a greater sensitivity to variations in testing conditions between occasions (cf. section on variance proportions above).

The ratio of covariances in part scores of different tests (c and f) to those of the same tests (b and e) indicates to what extent the two markers of the same factor measure the same kind of performance.[1/]

Similarity of performance between		Grades 4-6	Grades 7-9
Numerical tests	Class level	.79	1.00
	Pupil level	.84	.84
Perceptual tests	Class level	.75	.85
	Pupil level	.71	.69

At both levels similarity of performance is stronger between the numerical tests than between the perceptual tests. Only class level indices increase from elementary to upper grades. In numerical skills the class means for Addition and Multiplication have no specificity at all in grades 7-9, i.e., classes differ in overall numerical facility but not any longer in its components.

A corresponding index has been computed also for the similarity between numerical and perceptual facilities (row g in relation to the average of rows c and f).

Similarity of performance between		Grades 4-6	Grades 7-9
Numerical and Perceptual tests	Class level	.91	.89
	Pupil level	.77	.73

[1/] This measure is similar to a correlation corrected for attenuation with the exception that arithmetic instead of geometric means are used in the denominator.

The similarity is stronger at class than at pupil level. Earlier the difference between the two levels has found its expression in the factor analyses where N and P went together in the between-class analyses and were distinguished from each other in most within-class analyses.

Discussion and conclusions

The present study was initiated to serve as an empirical illustration to some of the methodological issues brought up by Cronbach in his paper on the choice of level for design and analysis in classroom research. The available data seem to have been well suited to exemplify some of the analytic procedures in treating multi-level data. They also illustrate the size of correlation ratios in an ordinary school situation under different grouping principles. The content of the tests used in the study ("primary mental abilities"), on the other hand, may not be well suited to register contextual effects of grouping or treatments varying between schools and classrooms. What is shown is mainly the sensitivity of different measures to pre-existing differences related to the grouping of pupils. Even so the study has become not only a methodological illustration but also resulted in some findings of substantive interest.

To begin with the decomposition of variance into parts for different levels, the analyses have made it possible to detect and interpret the following sources of variation.

Demographic differences between the areas from which the pupils are recruited are shown in district contributions to variance in all grades, and class contributions in the elementary grades. In general these variations are very small which may be due to two characteristics of the present data. They come mainly from rural and small town districts and

thus do not include, for instance, variations between big cities and suburban areas. Moreover they come from a country where, according to international comparisons (cf. Thorndike, 1973, p. 142), variations between schools in reading levels are small.

Self-selection effects are shown in the class contributions in grades 7 - 9. The choice between academic and non-academic programs in grades 7 - 8 gives rise to substantial class contributions in verbal, reasoning and numerical scores, i.e., in abilities that dominate in measures of IQ and scholastic performance. The further subdivision in vocational programs in grade 9 also registers in spatial tests. There is no reason to believe that the size of the self-selection effect found here is particularly small due to local or national charactersitics as above, but naturally it depends on the tracking system and the principles of class composition. For instance, the present Swedish system, with heterogeneous classes also in the upper grades of the compulsory school, should not originate such large effects.

Treatment effects are less likely to be found in ability than in achievement scores, but the numerical tests evidently function in both ways. Differences in emphasizing arithmetic skills are shown in district contributions in Addition and Multiplication, maybe also in the specificity for Addition vs. Multiplication in the elementary grades as found in the analysis of parallel forms.

Irregularities in test administration (cf. Gustafsson, Note 2) were used to explain some of the larger district and class contributions in speeded tests and the comparatively lower reliability of these tests at class level.

In this list of possible efiects one type is missing so far -- contextual effects caused by the interaction among pupils in the class. Such effects can be measured as differences between classes over and above what can be expected from the characteristics of the individual pupils. The present analysis does not permit detection of such effects, nor does the content of the tests make it plausible to find any.

From a substantive point of view the most interesting findings come from the factor analyses. Figure 1 presents an idealized picture of the results.

Insert Figure 1 about here

The main pattern is a hierarchical structure with four strata at individual level and three strata at collective level. The two upper strata are similar at both levels of analysis, the third is missing at class level where the "power" and "speed" factors are based directly on the tests, without the intervening level of primary factors. In some preliminary analyses of class means, however, where more than two factors were extracted in spite of their very small contributions to total variance, some of the "primary" factors could occasionally be traced. This might indicate that the similarity could extend also to the third stratum, though hardly detectable within the range of variation studied here. Furthermore all unrotated first-order analyses at individual level (3 x 6 analyses) have in their second component a contrast between the six "power" and the four "speed" tests. The same pattern holds for 17 out of 18 unrotated analyses at collective level. The similarity between the two levels of analysis thus is even more striking than the rotated

factors reported earlier indicate.

This similarity between individual and collective level suggests a sort of factorial isomorphism -- a hierarchical pattern repeating itself at varying levels of aggregation. On the other hand, class means are likely to vary with the principles on which aggregation is based and nominally identical variables may mean quite different things at different levels (cf. Cronbach, op. cit., p. 9.8 and 9.16-17). So there is reason to be suspicious about such ideas of factorial isomorphism between levels of analysis.

In the present study two different principles of class composition are represented: the neighborhood principle in grade 4 - 6 and the grouping-by-electives principle in grades 7 - 9. Both result in similar factor patterns at collective level but there is a great difference between the proportions of variance explained by different components. The following table summarizes the distributions of total variance (based on Table 4).

	Per cent of variance i grades	
	4-6	7-9
First component within	34	24
between	7	21
Second component within	18	17
between	3	3
Remainder within	35	31
between	3	4
Total	100	100

The change in aggregation principles affects the first components strongly and leaves the second and remaining components fairly unchanged. The growth of the first between-class component is very substantial -- from 7 to 21 percent. Even more interesting is that the second components, the contrasts between "power" and "speed" tests, are not affected at all. The conclusion so far is that the "power" vs. "speed" contrast has nothing to do with the change of grouping principles.

Further insight may be gained from the correlations between the separate "speed" tests at class and individual level (Table 9). Since the reliabilities have been shown to differ considerably between these tests (see Results III), the correlations have been corrected for attenuation.

Insert Table 9 about here

The four tests are here presented in a hypothetical order of complexity -- from simple perceptual scanning for identical numbers, over scanning in combination with size comparison, to addition and finally multiplication. The within-class correlations then form a simplex pattern (Guttman, 1954). The highest correlations are next to the diagonal and from there on correlations decrease in both rows and columns. This pattern holds both in elementary and upper grades and not only in the averages reported in the table but in all matrices as well, both for intact classes and for boys and girls separately. This ordering of perceptual and arithmetical operations with numbers, along a common scale of complexity, corresponds to what Guttman (1966, pp. 445-449) found when applying order analysis to Coombs' (1941) data on number ability. The regularity is even more striking

in the present study.

Between classes, on the other hand, no such ordering can be found. The correlations are fairly equal in size but with a surplus amount of commonality for Addition and Multiplication. Class level correlations are higher in the upper than in the elementary grades. The difference is reversed at individual level, which has to do with the size of between-class variation at the two stages. These differences are likely to have been absorbed by the first components.

In conclusion of this part of the discussion it seems evident that the change of grouping principles not only leaves the variance of the "power" vs. "speed" contrast unaffected but that also the processes operative in the "speed" tests are the same before and after the change. In contrast, different processes seem to be effective at individual and class level, and only the individual correlations can be explained by means of the order of complexity of the tasks.

From here on the present data are of little assistance to the interpretation of the factors at collective level. The following attempt needs support from measures of classroom process that are not available for these data.

In the elementary grades the distribution of variance is an expression of "natural" variation between unselected classes. Classes vary in general level due to differing demographic backgrounds of the pupils (the "power" factor) but also in their stress upon speed and accuracy of performance. When the grouping principle changes in grade 7 classes still show some "natural" variation in general level (the part associated with district variation, see Table 2), but in addition a strong self-selection sets in

according to academic performance and this is to varying degrees correlated with ability tests, mainly of the "power" type. Differences in emphasis upon speed and accuracy, on the other hand, remain at the same strength as before reorganization of classes. According to this interpretation the "speed" factor is more like a treatment or climate factor operating among the classes, maybe even mainly in test-like situations for skills that are not explicitly trained together in the school program as addition and multiplication are. The "power" factor is based on characteristics which the students bring with them when entering class.

So far the discussion has been based on the idealized picture in Figure 1. Some deviations from this picture should also be discussed. In grade 4 for intact classes, as well as for boys and girls separately, and in grade 5 for girls, the otherwise regular PMA pattern at individual level is broken and the tests sort themselves into the following groups. The parentheses indicate that the tests belong to the groups only in three of the four instances. Some small loadings appearing only once have been omitted.

Addition	Synonyms	Figure Series	Block Counting	Letter Grouping
Multiplication	Opposites	Metal Folding	(Metal Folding)	(Identical Numbers)
Identical Numbers				
Highest Number				

The overall picture is one of less structuring of abilities in the lowest grades and this is in agreement with the much discussed hypothesis of an increasing differentiation of abilities with age (cf. Guilford, 1967, pp. 412-417). Here the "full" PMA pattern does not appear until

grade 5 for boys and grade 6 for girls, i.e., at 12 and 13 years of age. From a general developmental viewpoint it would have been less surprising if the girls had been the first ones to show the full range of factors since they are earlier in most respects at this age level.

The merging of N and P factors in the youngest groups can easily be understood from the pattern of correlations between the tests at individual level. The average difference of correlations within vs. between the two categories of tests is a little larger in the matrices that give two factors than in the ones that merge them, but the difference is not great and results are in both cases consonant with the order-of-complexity hypothesis. To begin with I was surprised that the two factors merged in the younger and not in the older groups where an automatization of arithmetical operations was more likely to have taken place, making the two types of tasks less distinct from each other. But it is also possible that identification and comparison of numbers still are fairly complex tasks in the lower grades and therefore more similar to arithmetic, while automatization comes later and perhaps only in select groups of students.

The other irregularities were found in the reasoning and spatial tests. Figure Series and Metal Folding go together. Both are based on figural materials and from earlier studies (cf. e.g., French, 1965) it is known that some spatial tests can be solved by reasoning processes. An inspection of the correlations at individual level indicates that reasoning and spatial tests also may form simplex patterns. Table 10 shows average correlations in elementary and upper grades. Boys and girls have been kept separate because sex differences go in different directions

in Letter Grouping and the spatial tests.

Insert Table 10 about here

The simplex pattern holds for grades 4 - 6. In grades 7 - 9 the correlation between Letter Grouping and Block Counting deviates from the pattern. In this instance of a simplex pattern it is hardly meaningful to talk about an increasing order of complexity of the tasks. The ordering rather seems to express varying "mixes" of reasoning and visualization processes that go into the solution of the problems -- with identifying deviations from the principles of Letter Grouping at the reasoning end, and counting of partly concealed blocks in two-dimensional representations at the visualization end of the continuum. In some instances the intermediary tasks, Figure Series and Metal Folding, are close to each other and form a factor, leaving the extreme tests isolated at the ends. If only common variance had been analyzed, with a proper number of factors, and not as here total within-class variance, the extreme tests would possibly not have appeared as factors.

The analyses upon which these interpretations have been based deviate from traditional unrestricted factor analysis in a number of respects. Parallel analyses were performed at individual and class level. The analyses were based on scale-free covariances instead of correlations. The total variances at the two levels were analyzed instead of common variance. The number of factors was restricted in advance according to hypotheses.

Still another characteristic of the present design -- the parallel matrices from six grades -- could have been used for an additional restriction. A common hypothetical structure could have been tested against the data by means of Jöreskog's method for simultaneous factor analysis in several populations (Jöreskog, 1971). The finding of very similar factor structures from several independent analyses, using standard computational programs, indicates that this method might be fruitfully applied to the present data.

All analyses in the present study were done both for intact classes and for boys and girls separately. The separate analyses were done primarily as a safeguard against mean score sex differences influencing the structure of the covariance matrices -- increasing covariances between tests that have sex differences in the same direction and decreasing them in the opposite case. In general the results were rather similar in all sets of analyses.

This attempt to protect the analyses from effects of pooling data from boys and girls, however, brings up a central question in the discussion of multi-level analysis. Provided that the criteria of separation are not evident from theory and design, which degree of decomposition should be used in data analysis? Conventionally age (or grade) and sex variations are kept out of most analyses of educational data by means of grouping, or at least checked before aggregation. The reason for this is that these variables normally represent strong sources of pre-existing variation that can mask differences to be focussed in the particular study, or interact with them. School and class differences,

on the other hand, are not part of this convention of separating groups in the analysis. Even though sampling often is based on such units and not on the individual pupils, data are usually pooled before analysis. From the point of view of significance testing this practice is always doubtful. When treatments are given in groups and contextual effects are likely to appear Cronbach shows that it is misleading also from a theoretical point of view.

In a description of "natural" variation, on the other hand, pooling might be permissible since demographic variation between units is a legitimate part of the total variation to be described and analyzed. The aims of the study for which the present data were collected were primarily of this descriptive kind. Even in this case, however, the decomposition of variances and covariances has provided more distinctive information than the corresponding parts of the original pooled analysis. It seems to me that such decomposition is to be preferred even in a description of "natural" variation -- maybe followed by putting together the different sources of variation again to the full variation. However, since for economic reasons most collections of data permit only a limited number of grouping variables the problems remain: in what order of priority should different sources of variation be considered, when is it time to stop decomposition, and what level of analysis gives the most relevant information? These questions have to be answered in each study and preferably on the basis of theory rather than convention.

REFERENCE NOTES

1. Cronbach, L. J. Research on classrooms and schools: Formulation of questions, design, and analysis. Occasional Papers of the Stanford Evaluation Consortium, July 1976.

2. Gustafsson, J. E. A note on the importance of studying class class effects in aptitude-treatment interactions. Reports from the Institute of Education, Göteborg University, No. 52, 1976.

REFERENCES

Coombs, C. H. A factorial study of number ability. Psychometrika, 1941, 6, 161-189.

French, J. W. The relationship of problem-solving styles to the factor composition of tasks. Educational and Psychological Measurement, 1965, 25, 9-28.

French, J. W., Ekstrom, R. B., & Price, L. B. Kit of reference tests for cognitive factors. Princeton: Educational Testing Service, 1963.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Guttman, L. A new approach to factor analysis: the Radex. In: P. F. Lazarsfeld (Ed.), Mathematical thinking in the social sciences. Glencoe, Ill.: The Free Press, 1954.

Guttman, L. Order analysis of correlation matrices. In: R. B. Cattell (Ed.), Handbook of multivariate experimental psychology. Chicago: Rand McNally, 1966.

Härnqvist, K. Individuella differenser och skoldifferentiering. SOU 1960: 13, Stockholm.

Härnqvist, K. Canonical analyses of mental test profiles. Scandinavian Journal of Psychology, 1973, 14, 282-290.

Henrysson, S. Faktoranalys av Härnqvists anlagsprov. Nordisk Psykologi, 1965, 17, 11-20.

Jöreskog, K. G. Simultaneous factor analysis in several populations. Psychometrika, 1971, 36, 409-426.

Thorndike, R. L. Reading comprehension education in fifteen countries. New York: Wiley, 1973.

Winer, B. J. Statistical principles in experimental design. Second edition. New York: McGraw-Hill, 1971.

Table 1. Number of districts, classes, and pupils.

Grade:	4	5	6	7	8	9
Districts	5	5	5	5	5	5
Classes	29	26	28	20	21	27
Pupils, total	500	447	467	527	500	426
boys	274	232	228	255	250	210
girls	226	215	239	272	250	216
Average class size	17	17	17	26	24	16

Table 2. Relative contributions to total variance from different levels (in %). Intact classes.

	Districts						Classes within districts						Pupils within classes					
	4	5	6	7	8	9	4	5	6	7	8	9	4	5	6	7	8	9
Synonyms	2	2	1	3	7	5	13	9	11	32	23	27	85	88	88	65	70	68
Opposites	1	3	1	2	5	4	10	6	11	34	27	13	89	91	89	64	68	60
Letter Grouping	1	2	1	2	4	6	7	7	14	27	31	33	92	90	84	70	65	61
Figure Series	3	2	2	3	4	2	11	8	12	21	18	27	86	90	86	76	77	72
Metal Folding	1	2	1	2	3	5	10	7	12	15	17	26	90	91	87	83	79	69
Block Counting	0	1	2	0	2	1	9	5	7	12	11	27	91	94	91	88	87	72
Addition	3	1	2	7	9	3	9	11	13	18	22	24	88	88	85	74	68	73
Multiplication	3	2	3	6	6	5	14	12	14	20	27	29	84	86	83	75	66	65
Identical Numbers	3	2	1	2	5	2	10	8	23	11	24	27	87	89	76	88	71	71
Highest Number	2	2	3	10	18	4	10	17	27	13	13	21	88	81	69	77	70	75
Pattern Copying	5	6	5	5	1	1	19	7	13	15	20	22	76	88	82	80	79	76
Mark Making	2	3	6	6	6	6	19	15	26	27	28	34	78	82	68	68	66	61

Table 3. Examples of correlations and decomposed covariances (from grade 9, intact classes).

	Synonyms - Opposites	Synonyms - Metal Folding	Synonyms - Multiplication
Overall correlations	.800	.383	.440
Scale-free covariances			
between class means	.341	.262	.283
within classes	.459	.121	.157
Correlations			
between class means	.944	.822	.845
within classes	.722	.177	.236

Table 4. Distribution of total variance between principal components (in %). Intact classes.

		4	5	6	7	8	9
Within classes	comp. 1	32	37	33	24	26	22
	2	18	18	19	19	17	15
	3	9	9	8	9	7	8
	4	7	6	5	5	5	5
	5	6	5	5	5	5	5
	Remainder	16	14	14	14	12	14
	Subtotal	88	89	84	76	72	69
Between classes	comp. 1	6	5	9	18	20	24
	2	3	2	3	3	3	4
	Remainder	3	4	4	3	5	3
	Subtotal	12	11	16	24	28	31
	Total	100	100	100	100	100	100

Table 5. Within-class analyses: Factor loadings in five oblique factors. Intact classes. [1]

Grade	4					5					6					7					8					9				
Factor name	V	R	R/S	S	N/P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P
Factor number	2	5	3	4	1	1	5	3	2	4	1	4	3	2	5	3	5	2	1	4	1	5	2	3	4	3	5	4	1	2
Synonyms	86					92					94					77					80					77				
Opposites	86					87					83					72					75					68				
Letter Grouping		86					87					84					65					62					56			
Figure Series			79				61					63					76					80					77			
Metal Folding			67	32			28	60				63					61					62	-25				67			
Block Counting				89				92				91					88					88					78			
Addition					85			78					80					73					69					80		
Multiplication					78			86					87					77					72					74		
Identical Numbers		31			65				94					84					90					81					81	
Highest Number					78			26	61					57				31	52					63					65	

[1] Loadings above .24 are shown. Decimal points omitted.

Table 6. Within-class analyses: Second-order factor loadings.[1/] Intact classes.

First-order factor	4		First order factor	5		6		7		8		9	
	1	2		1	2	1	2	1	2	1	2	1	2
V	37	26	V	64	.	64	.	51	.	58	.	.	40
R	.	46	R	68	.	72	.	76	.	73	.	.	72
R/S	84	.	S	67	.	66	.	56	.	65	.	.	60
S	32	.	N	.	71	.	77	.	66	.	74	86	.
N/P	.	75	P	.	72	.	68	.	71	.	67	60	.

[1/] Loadings above .24 are shown. Decimal points omitted.

Table 7. Between-class analyses: Factor loadings[1/] in two oblique factors. Intact classes.

	4		5		6		7		8		9	
	1	2	1	2	1	2	1	2	1	2	1	2
Synonyms	33	.	.	26	.	(23)	52	.	50	.	54	.
Opposites	29	.	.	(24)	.	(22)	53	.	50	.	60	.
Letter Grouping	(18)	.	.	(17)	.	31	39	.	47	.	59	.
Figure Series	34	.	.	(24)	.	32	47	.	39	.	46	.
Metal Folding	27	.	.	(23)	.	33	39	.	45	.	45	28
Block Counting	(22)	.	.	(14)	.	(24)	30	.	33	.	40	26
Addition	.	31	28	.	34	.	.	41	.	39	46	.
Multiplication	.	37	30	.	32	.	.	37	35	31	55	.
Identical Numbers	.	29	28	.	47	.	.	32	.	29	44	.
Highest Number	.	(21)	35	.	47	.	.	47	.	51	37	-26

[1/] Loadings above .24 are shown. In addition, within parentheses, the highest loading for tests without a loading above .24. Decimal points omitted.

Table 8. Variances and covariances at class and pupil level in percent of overall variance in part scores. Intact classes.

	Grades 4-6			Grades 7-9		
	Classes	Pupils	Overall	Classes	Pupils	Overall
Numerical tests (Ad and Mu)						
a. Variance of part scores[1]	15	85	100	29	71	100
Covariance of part scores:						
b. in same test[2]	12	70	82	25	57	82
c. in different tests[1]	9	59	68	25	48	17
Perceptual speed tests (IN and HN)						
d. Variance of part scores[1]	19	81	100	26	74	100
Covariance of part scores:						
e. in same test[2]	13	65	78	17	55	72
f. in different tests[1]	10	46	56	13	38	51
Ad and Mu x IN and HN						
g. Covariance[3]	8	41	49	18	32	50

[1] Averages of 2 x 2 x 3 values. [2] Averages of 2 x 3 values. [3] Averages of 4 x 4 x 3 values.

Table 9. Average correlations between numerical and perceptual tests at different levels (corrected for attenuation). Intact classes.

	Grades 4-6			Grades 7-9		
Within classes	IN	HN	Ad	IN	HN	Ad
Identical Numbers (IN)						
Highest Number (HN)	.73			.69		
Addition (Ad)	.59	.71		.53	.66	
Multiplication (Mu)	.52	.62	.87	.50	.55	.84
Between classes						
HN	.72			.85		
Ad	.74	.72		.88	.87	
Mu	.73	.63	.79	.94	.81	1.00

Table 10. Average correlations between reasoning and spatial tests at different levels.

	Grades 4-6			Grades 7-9		
Boys	LG	FS	MF	LG	FS	MF
Letter Grouping (LG)						
Figure Series (FS)	.44			.46		
Metal Folding (MF)	.38	.51		.27	.42	
Block Counting (BC)	.33	.41	.53	.38	.38	.48
Girls						
FS	.49			.50		
MF	.42	.54		.38	.41	
BC	.41	.43	.51	.40	.39	.57

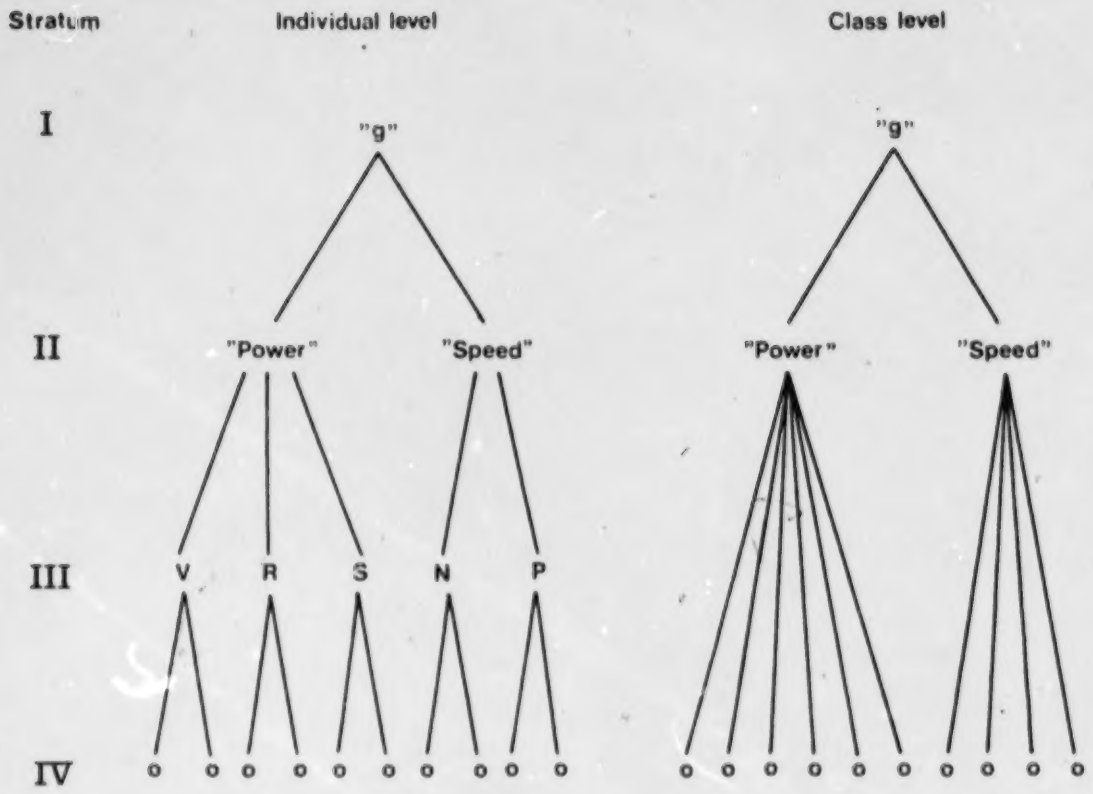

Figure 1. Factor structure at individual and class level

Appendix A1. Relative contributions to total variance from different levels (in %). Half-classes of boys.

	Districts						Classes within districts						Pupils within classes					
	4	5	6	7	8	9	4	5	6	7	8	9	4	5	6	7	8	9
Synonyms	3	6	1	4	8	8	18	12	14	33	24	29	79	82	85	63	68	63
Opposites	4	5	0	2	6	7	15	6	14	37	30	36	81	88	85	60	64	56
Letter Grouping	1	2	1	6	4	7	10	9	20	29	34	35	88	89	79	65	62	57
Figure Series	6	3	2	5	4	3	14	11	19	23	21	28	80	86	79	72	75	69
Metal Folding	0	5	2	3	2	10	15	11	16	16	21	19	85	84	82	81	77	70
Block Counting	1	3	5	2	1	3	11	7	11	19	16	21	88	89	84	80	83	76
Addition	5	2	12	16	12	3	10	14	20	17	23	26	85	84	79	67	66	71
Multiplication	6	4	2	13	6	4	18	12	24	16	28	34	77	84	74	70	66	61
Identical Numbers	10	4	3	8	15	2	17	10	27	11	15	25	74	86	70	81	70	74
Highest Number	4	2	4	12	23	3	11	18	32	15	13	22	84	80	64	73	64	75
Pattern Copying	1	11	5	10	1	5	22	8	17	15	23	25	77	80	77	75	76	69
Mark Making	5	5	5	10	7	7	21	21	36	24	23	39	74	74	58	66	70	54

Appendix A2. Relative contributions to total variance from different levels (in %). Half-classes of girls.

	Districts						Classes within districts						Pupils within classes					
	4	5	6	7	8	9	4	5	6	7	8	9	4	5	6	7	8	9
Synonyms	3	0	1	3	6	4	17	20	16	31	31	30	80	80	83	66	63	66
Opposites	0	2	2	2	5	3	12	15	15	29	31	40	88	83	84	68	64	57
Letter Grouping	3	4	3	1	4	6	16	12	20	24	32	33	81	84	78	75	63	61
Figure Series	3	2	2	1	6	3	16	9	18	21	19	30	81	89	80	78	75	67
Metal Folding	3	3	0	4	8	6	11	10	16	16	18	30	86	87	84	80	75	64
Block Counting	1	1	2	1	5	1	21	6	12	14	16	30	78	93	87	85	79	69
Addition	3	2	2	4	7	4	13	24	23	13	17	35	83	73	76	83	76	61
Multiplication	2	2	5	2	9	7	16	29	19	19	27	37	81	69	75	78	65	56
Identical Numbers	2	2	0	4	3	3	18	23	31	6	31	42	80	75	69	90	66	55
Highest Number	1	3	4	12	13	6	23	25	36	14	12	28	76	72	60	74	75	66
Pattern Copying	10	2	6	5	1	1	29	15	19	16	25	28	61	83	74	79	74	71
Mark Making	3	4	8	5	5	5	32	25	31	23	42	36	66	71	61	72	52	58

Appendix B. Distribution of total variance between principal components (in %). Half-classes with boys and girls.

		Boys						Girls					
		4	5	6	7	8	9	4	5	6	7	8	9
Within classes component	1	29	35	32	24	25	23	30	35	31	25	25	19
	2	17	15	17	18	16	14	16	17	16	18	16	14
	3	9	10	8	8	7	7	9	7	8	9	7	9
	4	7	6	5	5	5	6	6	5	6	6	6	5
	5	6	5	5	4	4	4	5	4	5	5	5	4
Remainder		14	14	11	12	11	13	15	13	11	15	11	11
Subtotal		82	85	78	71	70	67	81	81	77	78	70	62
Between classes component	1	8	6	11	21	21	23	8	10	13	15	22	30
	2	4	3	4	4	4	5	4	3	3	4	3	3
Remainder		6	6	7	4	5	5	7	6	7	3	5	5
Subtotal		18	15	22	29	30	33	19	19	23	22	30	38
Total		100	100	100	100	100	100	100	100	100	100	100	100

Appendix C1. Within-class analyses: Factor loadings in five oblique factors.[1] Half-classes with boys.

Grade	4					5					6					7					8					9				
Factor name	V	R	R/S	S	N/P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P	V	R/S	S	N/P	P
Factor number	2	5	4	3	1	1	3	5	2	4	1	4	3	2	5	2	4	3	1	5	1	5	3	4	2	3	4	5	1	2
Synonyms	80	.	.	.	.	89	.	.	.	.	92	.	.	.	.	76	.	.	.	.	79	.	.	.	.	75	.	.	.	.
Opposites	85	.	.	.	.	86	.	.	.	.	86	.	.	.	.	68	.	.	.	.	72	.	.	.	.	66	.	.	.	.
Letter Grouping	.	86	.	.	.	.	81	.	.	.	.	88	.	.	.	.	60	.	.	.	.	48	.	.	.		.	.	51	.	
Figure Series	.	.	79	.	.	.	69	.	.	.	.	42	.	.	.	.	79	.	.	.	.	84	.	.	.		49	36	.		
Metal Folding	.	.	38	58	.	.	39	52	.	.	.	59	-32			.	56	-27	.		.	52	-27	.			72	.	.	.	
Block Counting	.	.	.	90	.	.	86	.	.		.	88	.	.		.	82	.	.		.	89	.	.		.	82	.	.		
Addition	.	.	.	.	83	.	.	76	.		.	.	78	.		.	.	69	.		.	.	65	.		.	.	66	.	
Multiplication	.	.	.	.	72	.	.	79	.		.	.	81	.		.	.	79	.		.	.	72	.		.	.	63	.	
Identical Numbers	27	-26	.	57		.	.	.	93		.	.	.	74		.	.	.	90		.	.	.	79		.	.	.	82	
Highest Number	.	.	.	.	76	.	.	33	55		.	.	28	51		.	27	38	40		.	.	.	68		.	.	.	63	

[1] Loadings above .24 are shown. Decimal points omitted.

Appendix C2. Within-class analyses: Factor loadings [1] in five oblique factors. Half-classes with girls.

Grade	4					5					6					7					8					9				
Factor name	V	R	R/S	S	N/P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P	V	R	S	N	P
Factor number	2	5	3	4	1	2	5	3	4	1	1	4	3	2	5	4	5	2	1	3	3	4	2	1	5	1	4	2	3	5
Synonyms	85					81					90					76					74					78				
Opposites	82					82					78					76					73					69				
Letter Grouping		73				30	66				.61					.65	.25				.62					.57				
Figure Series			70				.90				.85					.78					.81					.77				
Metal Folding			77				.57	39				.57					.71					.73					.66				
Block Counting				83				.88				.91					.84					.78					.76				
Addition					80				.72			.74					.81					.74					.74				
Multiplication					76		-29		.61			.82					.74					.73					.64				
Identical Numbers		27			72				.72				.82						.92					.78					.68		
Highest Number					68				.76				30	47					.57				37	53					.53		

[1] Loadings above .24 are shown. Decimal points omitted.

Appendix D. Within-class analyses: Second-order factor loadings.[1/] Half-classes with boys and girls.

First-order factor	Boys 4		Girls 4		5	
	1	2	1	2	1	2
V	39	26	31	31	95	.
R	.	44	.	37	.	45
R/S	75	.	94	.	63	.
S	45	.	28	.	34	52
N/P	.	60	.	72	.	.

First-order factor	Boys 5		6		7		8	
	1	2	1	2	1	2	1	2
V	47	.	70	.	54	.	53	.
R	78	.	65	.	74	.	82	.
S	57	.	61	.	58	.	62	.
N	.	71	26	61	.	65	.	73
P	.	59	.	73	.	80	.	75

First-order factor	Girls 6		7		8		9	
	1	2	1	2	1	2	1	2
V	60	.	45	.	60	.	34	.
R	76	.	84	.	68	.	84	.
S	73	.	56	.	66	.	52	.
N	.	76	.	60	.	68	.	85
P	.	52	.	67	.	58	.	51

First-order factor	Boys 9	
	1	2
V	32	.
R/S	83	.
S	45	.
R/N	.	72
P	.	52

[1/] Loadings above .24 are shown. Decimal points omitted.

Appendix E. Between-class analyses: Factor loadings[1/] in two oblique factors. Half-classes with boys and girls.

	Boys												Girls											
	4		5		6		7		8		9		4		5		6		7		8		9	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Synonyms	.	38	29	.	.	(18)	58	.	46	.	50	.	34	.	.	36	.	(20)	51	.	59	.	31	29
Opposites	.	35	28	.	.	26	61	.	48	.	54	.	28	.	.	32	.	(22)	53	.	60	.	41	26
Letter Grouping	(15)		(13)	.	.	34	50	.	52	.	47	.	33	.	.	(16)	.	35	40	.	58	.	46	.
Figure Series	.	38	28	.	.	38	52	.	39	.	46	.	37	.	.	(22)	.	38	41	.	44	.	47	.
Metal Folding	.	35	32	.	.	35	39	.	46	.	52	.	25	.	.	30	.	34	43	.	46	.	60	.
Block Counting	.	(24)	(21)	.	.	35	32	.	34	.	41	.	26	.	.	(12)	.	26	32	.	41	.	55	.
Addition	29	.	.	28	39	.	.	43	.	45	.	48	.	35	45	.	43	.	.	35	32	.	.	55
Multiplication	42	.	.	30	45	.	.	41	25	36	.	53	.	34	48	.	29	.	.	33	47	.	.	50
Identical Numbers	46	.	.	28	51	.	.	45	.	48	.	47	.	36	48	.	53	.	.(21)		42	.	.	49
Highest Number	31	.	.	36	49	.	.	45	.	52	.	46	.	31	41	.	60	.	.	53	.	46	.	60

[1/] Loadings above .24 are shown. In addition, within parentheses, the highest loading for tests without a loading above .24. Decimal points omitted.

Appendix F. Correlations between factors.

	Grade	Within-class: Second-order	Between-class: First-order
Total	4	.31	.29
	5	.32	.27
	6	.19	.47
	7	.14	.50
	8	.22	.40
	9	.21	-.01
Boys	4	.31	.29
	5	.39	.21
	6	.23	.32
	7	.14	.56
	8	.23	.42
	9	.15	.48
Girls	4	.37	.24
	5	.35	.36
	6	.23	.50
	7	.18	.39
	8	.22	.27
	9	.21	.70

Reports from the Institute of Education, Göteborg University, 1976-

46. Gustafsson,Jan-Eric: Inconsistencies in aptitude-treatment interactions as a function of procedures in the studies and methods of analysis. Project MID 16, March 1976.

47. Gustafsson,Jan-Eric: Differential effects of imagery instructions on pupils with different abilities. Project MID 17, April 1976.

48. Ekholm, Mats: Social development in school. Summary and excerpts. Project SOS 23, May 1976.

49. Stangvik, Gunnar: Approaches to the analysis of learner-task interactions and some implications for the study of pedagogical processes. Project YP 7, January 1976.

50. Andrae, Annika: Non-graded instruction in small rural lower secondary schools. A presentation of the PANG-project. Paper read at the INTERSKOLA conference, July 1976, Project PANG 20, July 1976.

51. Patriksson, Göran: Attitudes toward olympic games of Swedish adolescents. Paper presented at the international congress of physical activity sciences in Quebec City 11-16 July 1976. September 1976.

52. Gustafsson, Jan-Eric: A note on the importance of studying class effects in aptitude-treatment interactions. Project MID 19, September 1976.

53. Gustafsson, Jan-Eric: Spatial ability and the suppression of visualization by reading. Project MID 20, September 1976.

54. Dahlgren,Lars Owe & Marton, Ference: Investigation into the learning and reaching of basic concepts in economics - a research project on higher education. Paper presented at the Congress of the European Association for Research and Development in Higher Education. August 30 - September 3, 1976, Laouvain la Neuve, Belgium, September 1976.

55. Marton, Ference: Skill as an aspect of knowledge. Some implications from research on students conceptions of central phenomena in their subjects. Paper presented at the Second International Conference on Improving University Teaching, July 13-16, 1976. Heidelberg, F.R. Germany, September 1976.

56. Andrae. Annika (Ed.): Non-graded instruction. Research organization and design. Administration and daily teaching experiences in small rural lower secondary schools. Experiences from the PANG-project. Paper read at the INTERSKOLA conference in Sveg, July 1976. Project PANG 23, August 1976.

57. Sandgren, Björn & Åsberg, Rodney: On cognition and social change. A report from a pilot study regarding the effect of schooling on cognitive growth and attitudes towards social change in Pakistan, October 1976.

58. Sandgren, Björn: Relation between cognition and social development. January 1977.

59. Entwistle, Noel: Changing approaches to research into personality and learning. February 1977.

60. Härnqvist, Kjell: Primary Mental Abilities at Collective and Individual Level. October 1977.